

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

**EP 0 853 126 A2**

(12)

**EUROPEAN PATENT APPLICATION**

(43) Date of publication:

15.07.1998 Bulletin 1998/29

(21) Application number: 98300130.6

(22) Date of filing: 08.01.1998

(51) Int Cl.<sup>6</sup>: **C12N 15/12, C07K 14/705,  
C07K 16/18, A61K 38/17,  
C12Q 1/68, G01N 33/53,  
G01N 33/68, C12N 5/10,  
A61K 48/00**

(84) Designated Contracting States:

**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC  
NL PT SE**

Designated Extension States:

**AL LT LV MK RO SI**

• Van Horn, Stephanie

King of Prussia, Pennsylvania 19406 (GB)

• Bergsma, Derk John

King of Prussia, Pennsylvania 19406 (GB)

• Mao, Joyce Yue

King of Prussia Pennsylvania 19406 (US)

(30) Priority: 10.01.1997 US 781250

(71) Applicant: **SMITHKLINE BEECHAM  
CORPORATION**

**Philadelphia Pennsylvania 19103 (US)**

(74) Representative: **Crump, Jullan Richard John et al**

**fJ Cleveland,**

**40-43 Chancery Lane**

**London WC2A 1JQ (GB)**

(72) Inventors:

• **Sathe, Ganesh**

**King of Prussia, Pennsylvania 19406 (GB)**

(54) **cDNA clone HE8CS41 that encodes a human 7-transmembrane receptor**

(57) HE8CS41 polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing HE8CS41 polypeptides and polynucleotides in the design of protocols for the treatment of infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkin-

son's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome, among others and diagnostic assays for such conditions.

**EP 0 853 126 A2**

## Description

## FIELD OF INVENTION

5 This invention relates to newly identified polynucleotides, polypeptides encoded by them and to the use of such polynucleotides and polypeptides, and to their production. More particularly, the polynucleotides and polypeptides of the present invention relate to G-protein coupled receptor, hereinafter referred to as HE8CS41. The invention also relates to inhibiting or activating the action of such polynucleotides and polypeptides.

## 10 BACKGROUND OF THE INVENTION

It is well established that many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second messengers, e.g., cAMP (Lefkowitz, Nature, 1991, 351:353-354). Herein these proteins are referred to as proteins participating in pathways with G-proteins or PPG proteins. Some examples of these proteins include the GPC receptors, such as those for adrenergic agents and dopamine (Kobilka, B.K., et al., Proc. Natl Acad. Sci., USA, 1987, 84:46-50; Kobilka, B.K., et al., Science, 1987, 238:650-656; Bunzow, J.R., et al., Nature, 1988, 336:783-787), G-proteins themselves, effector proteins, e.g., phospholipase C, adenylyl cyclase, and phosphodiesterase, and actuator proteins, e.g., protein kinase A and protein kinase C (Simon, M. I., et al., Science, 1991, 252:802-8).

20 For example, in one form of signal transduction, the effect of hormone binding is activation of the enzyme, adenylyl cyclase, inside the cell. Enzyme activation by hormones is dependent on the presence of the nucleotide GTP. GTP also influences hormone binding. A G-protein connects the hormone receptor to adenylyl cyclase. G-protein was shown to exchange GTP for bound GDP when activated by a hormone receptor. The GTP-carrying form then binds to activated adenylyl cyclase. Hydrolysis of GTP to GDP, catalyzed by the G-protein itself, returns the G-protein to its basal, inactive form. Thus, the G-protein serves a dual role, as an intermediate that relays the signal from receptor to effector, and as a clock that controls the duration of the signal.

The membrane protein gene superfamily of G-protein coupled receptors (otherwise known as 7TM receptors) has been characterized as having seven putative transmembrane domains. The domains are believed to represent transmembrane  $\alpha$ -helices connected by extracellular or cytoplasmic loops. G-protein coupled receptors include a wide range of biologically active receptors, such as hormone, viral, growth factor and neuroreceptors.

30 G-protein coupled receptors have been characterized as including these seven conserved hydrophobic stretches of about 20 to 30 amino acids, connecting at least eight divergent hydrophilic loops. The G-protein family of coupled receptors includes dopamine receptors which bind to neuroleptic drugs used for treating psychotic and neurological disorders. Other examples of members of this family include, but are not limited to, calcitonin, adrenergic, endothelin, cAMP, adenosine, muscarinic, acetylcholine, serotonin, histamine, thrombin, kinin, follicle stimulating hormone, opsins, endothelial differentiation gene-1, rhodopsins, odorant, and cytomegalovirus receptors.

Most G-protein coupled receptors have single conserved cysteine residues in each of the first two extracellular loops which form disulfide bonds that are believed to stabilize functional protein structure. The 7 transmembrane regions are designated as TM1, TM2, TM3, TM4, TM5, TM6, and TM7. TM3 has been implicated in signal transduction.

40 Phosphorylation and lipidation (palmitoylation or farnesylation) of cysteine residues can influence signal transduction of some G-protein coupled receptors. Most G-protein coupled receptors contain potential phosphorylation sites within the third cytoplasmic loop and/or the carboxy terminus. For several G-protein coupled receptors, such as the  $\beta$ -adrenoreceptor, phosphorylation by protein kinase A and/or specific receptor kinases mediates receptor desensitization.

45 For some receptors, the ligand binding sites of G-protein coupled receptors are believed to comprise hydrophilic sockets formed by several G-protein coupled receptor transmembrane domains, said socket being surrounded by hydrophobic residues of the G-protein coupled receptors. The hydrophilic side of each G-protein coupled receptor transmembrane helix is postulated to face inward and form polar ligand binding site. TM3 has been implicated in several G-protein coupled receptors as having a ligand binding site, such as the TM3 aspartate residue. TM5 serines, a TM6 asparagine and TM6 or TM7 phenylalanines or tyrosines are also implicated in ligand binding.

50 G-protein coupled receptors can be intracellularly coupled by heterotrimeric G-proteins to various intracellular enzymes, ion channels and transporters (see, Johnson et al., Endoc. Rev., 1989, 10:317-331). Different G-protein  $\alpha$ -subunits preferentially stimulate particular effectors to modulate various biological functions in a cell. Phosphorylation of cytoplasmic residues of G-protein coupled receptors have been identified as an important mechanism for the regulation of G-protein coupling of some G-protein coupled receptors. G-protein coupled receptors are found in numerous sites within a mammalian host.

Over the past 15 years, nearly 350 therapeutic agents targeting 7 transmembrane (7 TM) receptors have been successfully introduced onto the market.

This indicates that these receptors have an established, proven history as therapeutic targets. Clearly there is a need for identification and characterization of further receptors which can play a role in preventing, ameliorating or correcting dysfunctions or diseases, including, but not limited to, infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome.

## SUMMARY OF THE INVENTION

In one aspect, the invention relates to HE8CS41 polypeptides and recombinant materials and methods for their production. Another aspect of the invention relates to methods for using such HE8CS41 polypeptides and polynucleotides. Such uses include the treatment of infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome, among others. In still another aspect, the invention relates to methods to identify agonists and antagonists using the materials provided by the invention, and treating conditions associated with HE8CS41 imbalance with the identified compounds. Yet another aspect of the invention relates to diagnostic assays for detecting diseases associated with inappropriate HE8CS41 activity or levels.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the nucleotide and deduced amino acid sequence of human HE8CS41. SEQ ID NOS: 1 and 2.

## DESCRIPTION OF THE INVENTION

### Definitions

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"HE8CS41" refers generally to a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, or an allelic variant thereof.

"Receptor Activity" or "Biological Activity of the Receptor" refers to the metabolic or physiologic function of said HE8CS41 including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said HE8CS41.

"HE8CS41 polypeptides" refers to polypeptides with amino acid sequences sufficiently similar to HE8CS41 sequences, preferably exhibiting at least one biological activity of the receptor.

"HE8CS41 gene" refers to a polynucleotide having the nucleotide sequence set forth in SEQ ID NO:1 or allelic variants thereof and/or their complements.

"HE8CS41 polynucleotides" refers to polynucleotides containing a nucleotide sequence which encodes a HE8CS41 polypeptide or fragment thereof, or a nucleotide sequence which has at least 100 % identity to a nucleotide sequence encoding the polypeptide of SEQ ID NO:2 or the corresponding fragment thereof, or a nucleotide sequence which has sufficient identity to a nucleotide sequence contained in SEQ ID NO:1 to hybridize under conditions useable for amplification or for use as a probe or marker.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

"Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide" generally refers to any polyribonucleotide or polydeoxiribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes

DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications has been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993 and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol* (1990) 182:626-646 and Rattan *et al.*, "Protein Synthesis: Posttranslational Modifications and Aging", *Ann NY Acad Sci* (1992) 663 :48-62.

"Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Identity" is a measure of the identity of nucleotide sequences or amino acid sequences. In general, the sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published techniques. See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, 1988; BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, 1987; and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073. Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCS program package (Devereux, J., *et al.*, *Nucleic Acids Research* (1984) 12(1):387), BLASTP, BLASTN, FASTA (Atschul, S.F. *et al.*, *J Molec Biol* (1990) 215:403).

## Polypeptides of the Invention

The HE8CS41 polypeptides of the present invention include the polypeptide of SEQ ID NO:2 (in particular the mature polypeptide) as well as HE8CS41 polypeptides which have at least 60.6% identity to the polypeptide of SEQ ID NO:2 or the relevant portion and more preferably at least 80% identity, and still more preferably at least 90% identity, and even still more preferably at least 95% identity to SEQ ID NO: 2.

The HE8CS41 polypeptides may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

Biologically active fragments of the HE8CS41 polypeptides are also included in the invention. A fragment is a polypeptide having an amino acid sequence that entirely is the same as part, but not all, of the amino acid sequence of the aforementioned HE8CS41 polypeptides. As with HE8CS41 polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, and 101 to the end of HE8CS41 polypeptide. In this context "about" includes the particularly recited ranges larger or smaller by several, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes.

Preferred fragments include, for example, truncation polypeptides having the amino acid sequence of HE8CS41 polypeptides, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Also preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Biologically active fragments are those that mediate receptor activity, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those that are antigenic or immunogenic in an animal, especially in a human.

Thus, the polypeptides of the invention include polypeptides having an amino acid sequence at least 60.6% identical to that of SEQ ID NO:2 or fragments thereof with at least 60.6% identity to the corresponding fragment of SEQ ID NO:2. Preferably, all of these polypeptides retain the biological activity of the receptor, including antigenic activity. Included in this group are variants of the defined sequence and fragments. Preferred variants are those that vary from the referents by conservative amino acid substitutions -- i.e., those that substitute a residue with another of like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination.

The HE8CS41 polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

## Polynucleotides of the Invention

Another aspect of the invention relates to isolated polynucleotides which encode the HE8CS41 polypeptides and polynucleotides closely related thereto.

HE8CS41 of the invention is structurally related to other proteins of the G-protein coupled receptor, as shown by the results of sequencing the cDNA encoding human HE8CS41. The cDNA sequence contains an open reading frame encoding a protein of 370 amino acid residues with a deduced molecular weight of 41.89 kDa. HE8CS41 of Figure 1 (SEQ ID NO:2) has about 60.6% identity (using FASTA) in 292 amino acid residues with Chick probable G-protein coupled receptor 6H1 from T cells (Kaplan, M.H. et al, J. Immunol. 151: 628-636 (1993), Accession # P32250). Furthermore, HE8CS41 (SEQ ID NO:2) is 56.5 % identical to GCRT-human RB intron encoded putative GPCR over 313 amino acid residues (Toguchida J. et al, Genomics, 17: 535-543 (1993), Accession # P43657). Furthermore, HE8CS41 (SEQ ID NO:2) is 29.8 % identical to Melga Angio tensin II receptor over 312 amino acid residues (Murphy T.J. et al, Mol. Pharmacol. 44: 1-7 (1993), Accession # P33396). HE8CS41 gene of Figure 1 (SEQ ID NO:1) has about 100% identity (using BLAST) in 361 nucleotide residues with H.sapiens cDNA sequence, clone c-2bd07 (Accession # F07588, Tessier, A. et al, C. R. Acad. Sci. III Sci. VIE 318 (2), 263-272, 1995). Furthermore, HE8CS41 (SEQ ID NO: 1) is 99.71 % identical to homo sapiens cDNA clone 51646 polynucleotide over 350 nucleotide base residues (Accession

# H20663, Wilson, R. et al, WashU-Merck EST program, Unpublished, 1995).

One polynucleotide of the present invention encoding HEBCS41 may be obtained using standard cloning and screening, from a cDNA library derived from mRNA in cells of human fetal brain or human 8 week whole embryo using the expressed sequence tag (EST) analysis (Adams, M.D., *et al. Science* (1991) 252:1651-1656; Adams, M.D. *et al.*, *Nature*, (1992) 355:632-634; Adams, M.D., *et al.*, *Nature* (1995) 377 Supp:3-174). Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

Thus, the nucleotide sequence encoding HEBCS41 polypeptides may be identical over its entire length to the coding sequence in Figure 1 (SEQ ID NO:1), or may be a degenerate form of this nucleotide sequence encoding the polypeptide of SEQ ID NO:2, or may be highly identical to a nucleotide sequence that encodes the polypeptide of SEQ ID NO:2. Preferably, the polynucleotides of the invention contain a nucleotide sequence that is highly identical, at least 80% identical, with a nucleotide sequence encoding a HEBCS41 polypeptide, or at least 80% identical with the encoding nucleotide sequence set forth in Figure 1 (SEQ ID NO: 1), or at least 80% identical to a nucleotide sequence encoding the polypeptide of SEQ ID NO:2.

When the polynucleotides of the invention are used for the recombinant production of HEBCS41 polypeptide, the polynucleotide may include the coding sequence for the mature polypeptide or a fragment thereof, by itself; the coding sequence for the mature polypeptide or fragment in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc Natl Acad Sci USA* (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Among particularly preferred embodiments of the invention are polynucleotides encoding HEBCS41 polypeptides having the amino acid sequence of set out in Figure 1 (SEQ ID NO:2) and variants thereof.

Further preferred embodiments are polynucleotides encoding HEBCS41 variants that have the amino acid sequence of the HEBCS41 of Figure 1 (SEQ ID NO:2) in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acid residues are substituted, deleted or added, in any combination.

Further preferred embodiments of the invention are polynucleotides that are at least 80% identical over their entire length to a polynucleotide encoding the HEBCS41 polypeptide having the amino acid sequence set out in Figure 1 (SEQ ID NO:2), and polynucleotides which are complementary to such polynucleotides. In this regard, polynucleotides at least 85% identical over their entire length to the same are particularly preferred, and those with at least 90% are especially preferred. Furthermore, those with at least 97% are highly preferred and those with at least 98-99% are most highly preferred, with at least 99% being the most preferred.

The present invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences.

Polynucleotides of the invention, which are sufficiently identical to a nucleotide sequence contained in SEQ ID NO:1, may be used as hybridization probes for cDNA and genomic DNA, to isolate full-length cDNAs and genomic clones encoding HEBCS41 and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to the HEBCS41 gene. Such hybridization techniques are known to those of skill in the art. Typically these nucleotide sequences are 70% identical, preferably 80% identical, more preferably 90% identical to that of the referent. The probes generally will comprise at least 15 nucleotides. Preferably, such probes will have at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will range between 30 and 50 nucleotides.

The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to animal and human disease.

#### Vectors, Host Cells, Expression

The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, and host cells which are genetically engineered with vectors of the invention and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis *et al.*, *BASIC METHODS IN MOLECULAR BIOLOGY* (1986) and Sambrook *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring

Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989) such as calcium phosphate transfection, DEAE-dextran mediated transfection, transfection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used. Such systems include, among others, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL* (*supra*).

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the desired polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

If the HE8CS41 polypeptide is to be expressed for use in screening assays, generally, it is preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If HE8CS41 polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide; if produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

HE8CS41 polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

#### Diagnostic Assays

This invention also relates to the use of HE8CS41 polynucleotides for use as diagnostic reagents. Detection of a mutated form of HE8CS41 gene associated with a dysfunction will provide a diagnostic tool that can add to or define a diagnosis of a disease or susceptibility to a disease which results from under-expression, over-expression or altered expression of HE8CS41. Individuals carrying mutations in the HE8CS41 gene may be detected at the DNA level by a variety of techniques.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled HE8CS41 nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, e.g., Myers *et al.*, *Science* (1985) 230:1242. Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method. See Cotton *et al.*, *Proc Natl Acad Sci USA* (1985) 85: 4397-4401.

The diagnostic assays offer a process for diagnosing or determining a susceptibility to infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome through detection of mutation in the HE8CS41 gene by the methods described.

In addition, infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign

prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome, can be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of HE8CS41 polypeptide or HE8CS41 mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as an HE8CS41, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

#### Chromosome Assays

The nucleotide sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes). The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

#### Antibodies

The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them can also be used as immunogens to produce antibodies immunospecific for the HE8CS41 polypeptides. The term "immunospecific" means that the antibodies have substantial greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

Antibodies generated against the HE8CS41 polypeptides can be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., *Nature* (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today* (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, MONOCLONAL ANTIBODIES AND CANCER THERAPY, pp. 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.

Antibodies against HE8CS41 polypeptides may also be employed to treat infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome, among others.

#### Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with HE8CS41 polypeptide, or a fragment thereof, adequate to produce antibody and/or T cell immune response to protect said animal from infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as



Huntington's disease or Gilles de la Tourette's syndrome, among others. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which comprises, delivering HE8CS41 gene via a vector directing expression of HE8CS41 polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

Further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a HE8CS41 polypeptide wherein the composition comprises a HE8CS41 polypeptide or HE8CS41 gene. The vaccine formulation may further comprise a suitable carrier. Since HE8CS41 polypeptide may be broken down in the stomach, it is preferably administered parenterally (including subcutaneous, intramuscular, intravenous, intradermal etc. injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in-water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

#### Screening Assays

The HE8CS41 of the present invention may be employed in a screening process for compounds which bind the receptor and which activate (agonists) or inhibit activation of (antagonists) the receptor polypeptide of the present invention. Thus, polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See Coligan *et al.*, *Current Protocols in Immunology* 1(2):Chapter 5 (1991).

HE8CS41 proteins are ubiquitous in the mammalian host and are responsible for many biological functions, including many pathologies. Accordingly, it is desirous to find compounds and drugs which stimulate HE8CS41 on the one hand and which can inhibit the function of HE8CS41 on the other hand. In general, agonists are employed for therapeutic and prophylactic purposes for such conditions as infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome. Antagonists may be employed for a variety of therapeutic and prophylactic purposes for such conditions as infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome.

In general, such screening procedures involve producing appropriate cells which express the receptor polypeptide of the present invention on the surface thereof. Such cells include cells from mammals, yeast, *Drosophila* or *E. coli*. Cells expressing the receptor (or cell membrane containing the expressed receptor) are then contacted with a test compound to observe binding, or stimulation or inhibition of a functional response.

One screening technique includes the use of cells which express receptor of this invention (for example, transfected CHO cells) in a system which measures extracellular pH or intracellular calcium changes caused by receptor activation. In this technique, compounds may be contacted with cells expressing the receptor polypeptide of the present invention. A second messenger response, e.g., signal transduction, pH changes, or changes in calcium level, is then measured to determine whether the potential compound activates or inhibits the receptor.

Another method involves screening for receptor inhibitors by determining inhibition or stimulation of receptor-mediated cAMP and/or adenylate cyclase accumulation. Such a method involves transfecting a eukaryotic cell with the receptor of this invention to express the receptor on the cell surface. The cell is then exposed to potential antagonists in the presence of the receptor of this invention. The amount of cAMP accumulation is then measured. If the potential antagonist binds the receptor, and thus inhibits receptor binding, the levels of receptor-mediated cAMP, or adenylate cyclase, activity will be reduced or increased.

Another methods for detecting agonists or antagonists for the receptor of the present invention is the yeast based technology as described in U.S. Patent 5,482,835.

The assays may simply test binding of a candidate compound wherein adherence to the cells bearing the receptor is detected by means of a label directly or indirectly associated with the candidate compound or in an assay involving competition with a labeled competitor. Further, these assays may test whether the candidate compound results in a signal generated by activation of the receptor, using detection systems appropriate to the cells bearing the receptor at their surfaces. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Standard methods for conducting such screening assays are well understood in the art.

Examples of potential HE8CS41 antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligand of the HE8CS41, e.g., a fragment of the ligand, or small molecules which bind to the receptor but do not elicit a response, so that the activity of the receptor is prevented.

#### Prophylactic and Therapeutic Methods

This invention provides methods of treating an abnormal conditions related to both an excess of and insufficient amounts of HE8CS41 activity.

If the activity of HE8CS41 is in excess, several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as hereinabove described along with a pharmaceutically acceptable carrier in an amount effective to inhibit activation by blocking binding of ligands to the HE8CS41, or by inhibiting a second signal, and thereby alleviating the abnormal condition.

In another approach, soluble forms of HE8CS41 polypeptides still capable of binding the ligand in competition with endogenous HE8CS41 may be administered. Typical embodiments of such competitors comprise fragments of the HE8CS41 polypeptide.

In still another approach, expression of the gene encoding endogenous HE8CS41 can be inhibited using expression blocking techniques. Known such techniques involve the use of antisense sequences, either internally generated or separately administered. See, for example, O'Connor, *J Neurochem* (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Alternatively, oligonucleotides which form triple helices with the gene can be supplied. See, for example, Lee *et al.*, *Nucleic Acids Res* (1979) 6:3073; Cooney *et al.*, *Science* (1988) 241:456; Dervan *et al.*, *Science* (1991) 251:1360. These oligomers can be administered *per se* or the relevant oligomers can be expressed *in vivo*.

For treating abnormal conditions related to an under-expression of HE8CS41 and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound which activates HE8CS41, i.e., an agonist as described above, in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of HE8CS41 by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo*. For overview of gene therapy, see Chapter 20, *Gene Therapy and other Molecular Genetic-based Therapeutic Approaches*, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996).

#### Formulation and Administration

Peptides, such as the soluble form of HE8CS41 polypeptides, and agonists and antagonist peptides or small molecules, may be formulated in combination with a suitable pharmaceutical carrier. Such formulations comprise a therapeutically effective amount of the polypeptide or compound, and a pharmaceutically acceptable carrier or excipient. Such carriers include but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. Formulation should suit the mode of administration, and is well within the skill of the art. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

Preferred forms of systemic administration of the pharmaceutical compositions include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if properly formulated in enteric or encapsulated formulations, oral administration may also be possible. Administration of these compounds may also be topical and/or localized,

in the form of salves, pastes, gels and the like.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 µg/kg of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide *ex vivo*, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

#### Examples

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples illustrate, but do not limit the invention.

#### Example 1

A search of a random cDNA sequence database from Human Genome Sciences consisting of short sequences known as expressed sequence tags (EST) with 7-TM domains encoding cDNA sequences using BLAST algorithm disclosed an EST which was homologous to T cell receptors. In order to obtain the full length clone, a complete DNA sequence of the insert was deduced using automated DNA sequencing procedure. A map analysis of the DNA sequence using the GCG software indicated an open reading frame (ORF) consisting of 370 amino acid residues. Further analysis of the DNA sequence by FASTA and BLAST algorithms displayed the homology of this polypeptide sequence to the 7- transmembrane like G-protein coupled receptors. In addition, the hydrophobicity plot analysis using the lasergene protean software showed several features in common with G-protein linked receptors. Most prominent was the existence of seven hydrophobic regions of approximately 20-30 amino acids each, which are likely to represent membrane spanning domains providing the 7- transmembrane structural topology found among the G-protein linked superfamily of receptors. In order to confirm the identity of the clone further, PCR primers were designed using the nucleotide sequence of the open reading frame (ORF) and the DNA sequence was amplified from human fetal brain library. Correct size PCR bands were subcloned into PCRII vector from Stratagene and sequenced.

#### Example 2: Mammalian Cell Expression

The receptors of the present invention are expressed in either human embryonic kidney 293 (HEK293) cells or adherent dhfr CHO cells. To maximize receptor expression, all 5' and 3' untranslated regions (UTRs) are removed from the receptor cDNA prior to insertion into a CDN or pCDNA3 vector. The cells are transfected with individual receptor cDNAs by lipofectin and selected in the presence of 400 mg/ml G418. After 3 weeks of selection, individual clones are picked and expanded for further analysis. HEK293 or CHO cells transfected with the vector alone serve as negative controls. To isolate cell lines stably expressing the individual receptors, about 24 clones are typically selected and analyzed by Northern blot analysis. Receptor mRNAs are generally detectable in about 50% of the G418-resistant clones analyzed.

#### Example 3: Ligand Binding Assays

Ligand binding assays provide a direct method for ascertaining receptor function and are adaptable to a high throughput format. A bank of over 200 putative orphan receptor ligands has been assembled for initial screening. The bank comprises: transmitters, hormones and chemokines known to act via a human seven transmembrane (7TM) receptor; naturally occurring compounds which may be putative agonists for a human 7TM receptor, non-mammalian, biologically active peptides for which a mammalian counterpart has not yet been identified; and compounds not found in nature, but which activate 7TM receptors with unknown natural ligands. This bank is used to initially screen the receptor for known ligands. The minimal dissociation constant of a ligand for its receptor in order to configure a reliable radioligand binding assay is about 10 nM.

The purified ligand for a receptor is radiolabeled to high specific activity (50-1000 Ci/mmol) for binding studies. A determination is then made that the process of radiolabeling does not diminish the activity of the ligand towards its receptor. Assay conditions for buffers, ions, pH and other modulators such as nucleotides are optimized to establish

a workable signal to noise ratio for both membrane and whole cell receptor sources. For these assays, specific receptor binding is defined as total associated radioactivity minus the radioactivity measured in the presence of an excess of unlabeled competing ligand. Where possible, more than one competing ligand is used to define residual nonspecific binding. About 50% specific binding is considered optimal.

The relevance of the radioligand binding site as a receptor is defined by the following criteria. These criteria are derived from the definition of a receptor as a moiety that specifically recognizes the appropriate agonist and antagonist ligands through binding interactions and initiates a defined biological response. The criteria include: (1) the kinetics of radioligand binding approximating the kinetics of the agent's effect on the biological response; (2) the specific binding of increasing concentrations of radioligand becomes saturated thus reflecting a finite receptor population; (3) the concentration range over which the radioligand binds parallels the concentration range over which the ligand activates or inhibits a biological response; and (4) the receptor sites occupied by the radioligand demonstrate the appropriate pharmacological specificity as the biological response.

#### Example 4: Functional Assay in *Xenopus* Oocytes

Capped RNA transcripts from linearized plasmid templates encoding the receptor cDNAs of the invention are synthesized in vitro with RNA polymerases in accordance with standard procedures. In vitro transcripts are suspended in water at a final concentration of 0.2 mg/ml. Ovarian lobes are removed from adult female toads, Stage V defolliculated oocytes are obtained, and RNA transcripts (10 ng/oocyte) are injected in a 50 nl bolus using a Drummond microinjection apparatus. Two electrode voltage clamps are used to measure the currents from individual *Xenopus* oocytes. Recordings are made in Ca<sup>2+</sup> free Barth's medium at room temperature.

#### Example 5: Microphysiometric Assays

Activation of a wide variety of secondary messenger systems results in extrusion of small amounts of acid from a cell. The acid formed is largely as a result of the increased metabolic activity required to fuel the intracellular signaling process. The pH changes in the media surrounding the cell are very small but are detectable by the CYTOSENSOR microphysiometer (Molecular Devices Ltd., Menlo Park, CA). The CYTOSENSOR is thus capable of detecting the activation of a receptor which is coupled to an energy utilizing intracellular signaling pathway such as the G-protein coupled receptor of the present invention.

#### Example 6: Extract/Cell Supernatant Screening

A large number of mammalian peptides exist for which there remains, as yet, no mammalian equivalent. Thus, active ligands for this receptor may not be included within the ligands banks as identified to date. Accordingly, the TM receptor of the invention is also screened against tissue extracts to identify natural ligands.

#### Example 7: Calcium Functional Assay

7TM receptors which are expressed in HEK 293 cells have been shown to be coupled functionally to activation of PLC and calcium mobilization. Basal calcium levels in the HEK 293 cells in receptor-transfected or vector control cells were observed to be in the normal, 100 nm to 200 nm, range. HEK 293 cells expressing recombinant receptors are loaded with fura 2 and in a single day > 150 selected ligands are evaluated for agonist induced calcium mobilization. Agonists presenting a calcium transient are tested in vector control cells to determine if the response is unique to the transfected receptor cells. When a unique agonist-induced response is identified, the response will be reproduced in a separate group of cells and then pharmacologically characterized with concentration response curves for the effective and related ligands.

## SEQUENCE LISTING

5

## (1) GENERAL INFORMATION

10

(i) APPLICANT: SmithKline Beecham Corporation

15

(ii) TITLE OF THE INVENTION: cDNA CLONE HE8CS41 THAT ENCODES A NOVEL  
7-TRANSMEMBRANE RECEPTOR

20

(iii) NUMBER OF SEQUENCES: 2

25

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: F J Cleveland &amp; Company

(B) STREET: 40/43 Chancery Lane

(C) CITY: London

(D) STATE:

(E) COUNTRY: United Kingdom

30

(F) POST CODE: WC2A 1JQ

35

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette

(B) COMPUTER IBM Compatible

(C) OPERATING SYSTEM: DOS

40

(D) SOFTWARE: FastSEQ for Windows Version 2.0

45

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: Unknown

(B) FILING DATE: 10-JAN-1997

(C) CLASSIFICATION:

50

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

55

## (viii) ATTORNEY/AGENT INFORMATION

- (A) NAME: CRUMP, Julian Richard John  
 (B) GENERAL AUTHORISATION NUMBER 37127  
 (C) REFERENCE/DOCKET NUMBER ATG50043

## (ix) TELECOMMUNICATION INFORMATION

- (A) TELEPHONE: +44 171 405 5875  
 (B) TELEFAX: +44 171 831 0749  
 (C) TELEX:

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 1640 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION SEQ ID NO: 1:

```

GGCAGAGAC CCTGCCCAAT GGTAACAGTT GTTGCCAAGA AAAGAAAAAA AGAAAAAAA    60
AAGACCCAAT CACCCATTGG AGGAAATCTC TAACTACAGT GCCTCCAGGA AACTCATTGG    120
TTATTCCAGT CTTAGATAAA ATTGGTGCCA GGATTGGTT CGAACTAATG TGGAGGAAAA    180
TATTTCTAC  CGGTCCATAG TGT CAGAGTG GTGAACCCCT GCAGCCAGCA GGCCTCCTGA    240
AAAAAAGTC  CATGGGTGAC AGAAGATTCA TTGACTTCCA ATTCCAAGAT TCAAATTCAA    300
GCCTCAGACC CAGGTTGGGC AATGCTACTG CCAATAATAC TTGCATTGTT GATGATTCTT    360
TCAAGTATAA TCTCAATGGT GCTGTCTACA GTGTGTATT CATCTGGGT CTGATAACCA    420
ACAGTGTCTC TCTGTTTGT CTTCTGTTTCC GCATGAAAAT GAGAAGTGAG ACTGCTATTT    480
TTATACCAA  TCTAGCTGTC TCTGATTGTC TTTTGTCTG TACACTACCT TTTAAATAT    540
TTTACAACTT CAACGCCAC TGGCCTTTTG GTGACACCCT CTGCAAGATC TCTGGAACTG    600
CATTCTTAC  CAACATCTAT GGGAGCATGC TCTTCTCAC CTGTATTAGT GTGGATCGTT    660

```

TCCTGGCCAT TGTCTATCCT TTTCGATCTC GTACTATTAG GACTAGGAGG AATTCTGCCA 720  
 TTGTGTGTGC TGGTGTCTGG ATCCTAGTCC TCAGTGGGG TATTTAGCC TCTTTGTTTT 780  
 5 CCACCACTAA TGTCAACAAT GCAACCACCA CCTGCTTTGA AGGCTTCTCC AAACGTGTCT 840  
 GGAAGACTTA TTTATCCAAG ATCACAATAT TTATTGAAGT TGTGGGTTT ATCATTCTCTC 900  
 TAATATTGAA TGTCTCTTGC TCTTCTGTGG TGCTGAGAAC TCTTCGCAAG CCTGCTACTC 960  
 10 TGTCTCAAAT TGGGACCAAT AAGAAAAAAG TACTGAAAAT GATCACAGTA CATATGGCAG 1020  
 TCTTTGTGGT ATGCTTTGTA CCCTACAAC CTGTCTCTT CTTGTATGCC CTGGTGGCT 1080  
 CCCAAGCTAT TACTAATTGC TTTTGGAAA GATTGCAAA GATCATGTAC CCAATCACCT 1140  
 TGTGCCTTGC AACTCTGAAC TGTTGTTTTG ACCCTTCAT CTATTACTTC ACCCTGAAT 1200  
 15 CCTTTCAGAA GTCCTTCTAC ATCAATGCCC ACATCAGAAT GGAGTCCCTG TTTAAGACTG 1260  
 AAACACCTTT GACCACAAAG CCTTCCCTTC CAGCTATTCA AGAGGAAGTG AGTGATCAAA 1320  
 CAACAATAAA TGGTGGTGAA TTAATGCTAG AATCCACCT TTAGGTATGA GAAATGTGT 1380  
 20 CAGGTCCAGA TATGGTTTCT CCTATAATTT TTCTATGCT ATAACTAAA GATTTGAAGC 1440  
 TAATGATACT GAGAATAATG CACCAAATCC AGTCAGATAC ATTTGTTTGA AGGTATACTG 1500  
 TAGAGTTTTT ATTGCTGTTT TGTTAGTAA TTATAGGTCA AATCTAATTA CAACAACCAA 1560  
 25 GATGGATTGC CAACTCTTC TGCTTGGTTG GAATTCATT GTATGCATT ATCCAGGTGG 1620  
 CCAGTGGCAT TTGATAATAT 1640

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 370 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Gly Asp Arg Arg Phe Ile Asp Phe Gln Phe Gln Asp Ser Asn Ser  
 1 5 10 15  
 50 Ser Leu Arg Pro Arg Leu Gly Asn Ala Thr Ala Asn Asn Thr Cys Ile  
 20 25 30  
 Val Asp Asp Ser Phe Lys Tyr Asn Leu Asn Gly Ala Val Tyr Ser Val  
 35 40 45  
 55 Val Phe Ile Leu Gly Leu Ile Thr Asn Ser Val Ser Leu Phe Val Phe

	50	55	60
	Cys Phe Arg Met Lys	Met Arg Ser Glu Thr Ala	Ile Phe Ile Thr Asn
5	65	70	75 80
	Leu Ala Val Ser Asp	Leu Leu Phe Val Cys Thr	Leu Pro Phe Lys Ile
	85	90	95
10	Phe Tyr Asn Phe Asn Arg His Trp	Pro Phe Gly Asp Thr	Leu Cys Lys
	100	105	110
	Ile Ser Gly Thr Ala Phe	Leu Thr Asn Ile Tyr Gly	Ser Met Leu Phe
	115	120	125
15	Leu Thr Cys Ile Ser Val Asp Arg	Phe Leu Ala Ile Val Tyr	Pro Phe
	130	135	140
	Arg Ser Arg Thr Ile Arg Thr Arg Arg	Asn Ser Ala Ile Val Cys Ala	
20	145	150	155 160
	Gly Val Trp Ile Leu Val Leu Ser Gly Gly	Ile Ser Ala Ser Leu Phe	
	165	170	175
25	Ser Thr Thr Asn Val Asn Asn Ala Thr Thr Thr	Cys Phe Glu Gly Phe	
	180	185	190
	Ser Lys Arg Val Trp Lys Thr Tyr Leu Ser Lys	Ile Thr Ile Phe Ile	
	195	200	205
30	Glu Val Val Gly Phe Ile Ile Pro Leu Ile Leu Asn Val	Ser Cys Ser	
	210	215	220
	Ser Val Val Leu Arg Thr Leu Arg Lys Pro Ala Thr Leu	Ser Gln Ile	
35	225	230	235 240
	Gly Thr Asn Lys Lys Lys Val Leu Lys Met Ile Thr Val His Met Ala		
	245	250	255
40	Val Phe Val Val Cys Phe Val Pro Tyr Asn Ser Val Leu Phe Leu Tyr		
	260	265	270
	Ala Leu Val Arg Ser Gln Ala Ile Thr Asn Cys Phe Leu Glu Arg Phe		
	275	280	285
45	Ala Lys Ile Met Tyr Pro Ile Thr Leu Cys Leu Ala Thr Leu Asn Cys		
	290	295	300
	Cys Phe Asp Pro Phe Ile Tyr Tyr Phe Thr Leu Glu Ser Phe Gln Lys		
50	305	310	315 320
	Ser Phe Tyr Ile Asn Ala His Ile Arg Met Glu Ser Leu Phe Lys Thr		
	325	330	335
55	Glu Thr Pro Leu Thr Thr Lys Pro Ser Leu Pro Ala Ile Gln Glu Glu		
	340	345	350



Val Ser Asp Gln Thr Thr Asn Asn Gly Gly Glu Leu Met Leu Glu Ser

355

360

365

Thr Phe

370

## Claims

1. An isolated polynucleotide comprising a nucleotide sequence that has at least 80% identity to a nucleotide sequence encoding the polypeptide of SEQ ID NO:2 or the corresponding fragment thereof; or a nucleotide sequence complementary to said nucleotide sequence.
2. The polynucleotide of claim 1 which is DNA or RNA.
3. The polynucleotide of claim 1 wherein said nucleotide sequence is at least 80% identical to that contained in SEQ ID NO: 1.
4. The polynucleotide of claim 3 wherein said nucleotide sequence is contained in SEQ ID NO: 1.
5. The polynucleotide of claim 1 wherein said encoding nucleotide sequence encodes the polypeptide of SEQ ID NO:2 or a fragment thereof.
6. A polynucleotide probe or primer comprising at least 15 contiguous nucleotides of the polynucleotide of claim 3.
7. A DNA or RNA molecule comprising an expression system wherein said expression system is capable of producing a HE8CS41 or a fragment thereof having at least 80% identity with a nucleotide sequence encoding the polypeptide of SEQ ID NO:2 or said fragment when said expression system is present in a compatible host cell.
8. A host cell comprising the expression system of claim 7.
9. A process for producing a HE8CS41 polypeptide or fragment comprising culturing a host of claim 8 and under conditions sufficient for the production of said polypeptide or fragment.
10. The process of claim 9 wherein said polypeptide or fragment is expressed at the surface of said cell.
11. Cells produced by the process of claim 10.
12. The process of claim 9 which further includes recovering the polypeptide or fragment from the culture.
13. A process for producing a cell which produces a HE8CS41 polypeptide or a fragment thereof comprising transforming or transfecting a host cell with the expression system of claim 7 such that the host cell, under appropriate culture conditions, produces a HE8CS41 polypeptide or fragment.
14. A HE8CS41 polypeptide or a fragment thereof comprising an amino acid sequence which is at least 80% identical to the amino acid sequence contained in SEQ ID NO:2.
15. The polypeptide of claim 14 which comprises the amino acid sequence of SEQ ID NO:2, or a fragment thereof.
16. A HE8CS41 polypeptide or fragment prepared by the method of claim 12.
17. An antibody immunospecific for the HE8CS41 polypeptide of claim 14.
18. A method for the treatment of a subject in need of enhanced HE8CS41 activity comprising:

- (a) administering to the subject a therapeutically effective amount of an agonist to said receptor; and/or
- (b) providing to the subject HE8CS41 polynucleotide in a form so as to effect production of said receptor activity *in vivo*.

19. A method for the treatment of a subject having need to inhibit HE8CS41 activity comprising:

- (a) administering to the subject a therapeutically effective amount of an antagonist to said receptor; and/or
- (b) administering to the subject a nucleic acid molecule that inhibits the expression of the nucleotide sequence encoding said receptor; and/or
- (c) administering to the subject a therapeutically effective amount of a polypeptide that competes with said receptor for its ligand.

20. A process for diagnosing a disease or a susceptibility to a disease in a subject related to expression or activity of HE8CS41 in a subject comprising:

- (a) determining the presence or absence of a mutation in the nucleotide sequence encoding said HE8CS41 in the genome of said subject; and/or
- (b) analyzing for the presence or amount of the HE8CS41 expression in a sample derived from said subject.

21. A method for identifying compounds which bind to HE8CS41 comprising:

- (a) contacting cells of claim 11 with a candidate compound; and
- (b) assessing the ability of said candidate compound to bind to said cells.

22. The method of claim 21 which further includes determining whether the candidate compound effects a signal generated by activation of the HE8CS41 polypeptide at the surface of the cell, wherein a candidate compound which effects production of said signal is identified as an agonist.

23. An agonist identified by the method of claim 22.

24. The method of claim 21 which further includes contacting said cell with a known agonist for said HE8CS41; and determining whether the signal generated by said agonist is diminished in the presence of said candidate compound, wherein a candidate compound which effects a diminution in said signal is identified as an antagonist for said HE8CS41.

25. An antagonist identified by the method of claim 24.

FIGURE 1. Nucleotide and Amino Acid sequence of HE8CS41 (SEQ ID NOS: 1 and 2, respectively.)

1	GGCAOGAGACCCCTGCCAATGGTAAACAGTTGTTGCCAAGAAAAGAAAAAGAAAAAAA	60
61	AAGACCCAAT CACCCATTGGAGGAAAT CTCTAACTACAGTGCCTCCAGGAAACTCATTGG	120
121	TTATTCCAGTCTTAGATAAAATTGGTGCCAGGATTTGGTTGGAACTAATGTGGAGGAAAA	180
181	TATTTCTACCGGTCCATAGTGTGAGAGTGGTGAAACCCCTGCAGCCAGCAGGCCTCCTGA	240
241	AAAAAAGTCCATGGGTGACAGAAGATT CATTGACTTCCAATTCCAAGATTCAAATCAA	300
163	M G D R R F I D F Q F Q D S N S S	182
301	GCCTCAGACCCAGGTTGGGCAATGCTACTGCCAATAATACTTGCAATTGTTGATGATTCT	360
183	L R P R L G N A T A N N T C I V D D S F	202
361	TCAAGTAT AATCTCAATGGTGCTGTCTACAGTGTGTTGTT CATCTTGGGTCTGATAACCA	420
203	K Y N L N G A V Y S V V F I L G L I T N	222
421	ACAGTGTCTCTCTGTTTGTCTTCTGTTTCCGCATGAAAATGAGAAGTGAGACTGCTATTT	480
223	S V S L F V F C F R M K M R S E T A I F	242
481	TTATCACCATCTAGCTGTCTCTGATTGCTTTTTGTCTGTACACTACCTTTAAAATAT	540
243	I T N L A V S D L L F V C T L P F K I F	262
541	TTTACAACTTCAACCGCACTGGCCTTTTGGTGACACCTCTGCAAGATCTCTGGAACTG	600
263	Y N F N R H W P F G D T L C K I S G T A	282
601	CATTCTTACCAACATCTATGGGAGCATGCTCTTTCTCACCTGTATTAGTGTGGATCGTT	660
283	F L T N I Y G S M L F L T C I S V D R F	302
661	TCCTGGCCATTGTCTATCCTTTTCGATCTCGTACTATTAGGACTAGGAGGAATTCTGCCA	720
303	L A I V Y P F R S R T I R T R R N S A I	322
721	TTGTGTGTGCTGGTGTCTGGATCCTAGTCTCAGTGGGGTATTT CAGCCTCTTTGTTTT	780
323	V C A G V W I L V L S G G I S A S L F S	342
781	CCACCACTAATGTCAACATGCAACCAACCTGCTTTGAAGGCTTCTCCAAACGTGTCT	840
343	T T N V N N A T T T C F E G F S K R V W	362
841	GGAAGACTATTATCCAAGATCACAATATTTATTGAAGTTGTTGGGTTTATCATTCTC	900
363	K T Y L S K I T I F I E V V G F I I P L	382
901	TAATATTGAATGTCTCTTGCTCTTCTGTGGTGCTGAGAACTCTTGGCAAGCCTGCTACTC	960
383	I L N V S C S S V V L R T L R K P A T L	402
961	TGTCTCAAATTGGGACCAAT AAGAAAAAGTACTGAAAATGATCACAGTACATATGGCAG	1020
403	S Q I G T N K K K V L K M I T V H M A V	422
1021	TCCTTGTGGTATGCTTTGTACCTACAACTCTGTCTCTCTGTATGCCCTGGTGCGCT	1080
423	F V V C F V P Y N S V L F L Y A L V R S	442
1081	CCCAAGCTATTACTAATTGCTTTTGGAAAGATTTGCAAAGATCATGTACCAATCACCT	1140
443	Q A I T N C F L E R F A K I M Y P I T L	462
1141	TGTGCCTTGCAACTCTGAACGTGTGTTTGACCCTTTCATCTATTACTTCAACCTTGAAT	1200
463	C L A T L N C C F D P F I Y Y F T L E S	482
1201	CCTTTCAGAAGTCTTCTACATCAATGCCACATCAGAATGGAGTCCCTGTTTAAGACTG	1260
483	F Q K S F Y I N A H I R M E S L F K T E	502
1261	AAACACCTTTGACCACAAAGCCTTCCCTTCCAGCTATTCAAGAGGAAGTGAGTGATCAA	1320
503	T P L T T K P S L P A I Q E E V S D Q T	522

1321	CAACAAAT AATGGTGGTGAATT AATGCTAGAAT CCACCTTTTAGGTATGAGAAATGTGTT	1380
523	T N N G G E L M L E S T F *	542
1381	CAGGT CCAGATATGGTTT CT CCTAT AATTTTT CCTATGCTAT AAAC TAAAGATTGAAGC	1440
1441	TAATGATACTGAGAATAATGCACCAAAT CCAGT CAGATACATTTGTTTGAAGGTATCTG	1500
1501	TAGAGTTTTTATTGCTGTTTGT CAGTAATTATAGGT CAAATCTAATTACAACAACCA	1560
1561	GATGGATTGCCAAACTCTTCTGCTTGGTTGGAATTCATTGTATCGCATTATCCAGGTGG	1620
1621	CCAGTGGCATTGATAATAT	1640